DOCUMENT RESUME

ED 091 425 TM 003 639

AUTHOR Braverman, Barbara
TITLE The Applicability of Item Sampling to the Monitoring of Resourcefulness.
PUB DATE [Apr 74]
NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Achievement; Alternative Schools; Classroom Observation Techniques; Criterion Referenced Tests; *Curriculum Evaluation; *Item Sampling; Junior High Schools; *Measurement Techniques; Resource Materials; Student Development
IDENTIFIERS *Unobtrusive Measures

ABSTRACT

The usefulness of item sampling with unobtrusive measures and observational data is demonstrated. Particular emphasis is given to the question: Is item sampling appropriate for estimating the differences between means as well as the means themselves? A descriptive analysis is made of one aspect of a junior high school curriculum. Twenty four students were administered nine items from the "Resourcefulness Test" on three occasions. Item samples were then generated a posteriori from each 24 x 9 matrix, using balanced incomplete block design. Difference scores were examined. A description of student performance on each item across occasions constituted the curriculum evaluation, or monitoring part of the study. The results show that item sampling appears to be an effective technique for monitoring unobtrusive measures and observational data. (Author/BB)

ED091425

TM 003 639

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE APPLICABILITY OF ITEM SAMPLING TO THE MONITORING OF RESOURCEFULNESS

A PAPER PRESENTED AT THE AMERICAN EDUCATIONAL RESEARCH ASSOICATION CHICAGO, ILLINOIS, APRIL 17,1974

Dr. Barbara Braverman
National Technical Institute
for the Deaf
Rochester, New York, 14623

INTRODUCTION

An alternative approach to curriculum evaluation has a particular advantage in "open education" settings where students are participating in many different educational experiences. It has further meaning when individual differences are of less concern than group differences in the evaluation of curricular objectives.

The purpose of this paper is both methodological and descriptive. An adaptation of achievement monitoring was employed to determine the extent to which one of the objectives of a learner-centered program at the junior high school level had been satisfied. The applicability of the item sampling approach to this problem was explored.

The model used for the evaluation is analogous to the one-sample Time-Series Design (Campbell and Stanley, 1963). This study was concerned with behaviorally defined cognitive and affective measures, however, rather than traditional achievement items.

Particular emphasis is given to the question:

Is item sampling appropriate for estimating the differences between means as well as the means themselves?

An associated purpose of the paper is an interpretation of the data as a partial evaluation of the curriculum at an "alternative" junior high school in Rochester, New York. One of the school's objectives is to develop "resourcefulness" in its students. That objective has been operationalized in this study and a "resourcefulness test" has been developed to

measure the objective. A statement about changes in "resourcefulness" of youngsters in this setting has implications for program effectiveness. Although such differences cannot be attributed exclusively to the effect of the school's program on pupil development, the monitoring technique's sensitivity to changes over time is important to consider.

Estimates of mean score and standard deviation by item sampling have been shown to be fruitful by Lord (1962) Plumlee (1964), Knapp (1968), Shoemaker (1971) and others. In each preceeding case cognitive data was used. The usefulness of item sampling with attitude scales has been considered in several studies (Bursack and Cook, 1970; Sirotnik, 1970; Pugh, 1971; Shoemaker, 1971; Peterson and Anderson, 1971). The evidence indicates that item sampling techniques can be applied to affective measures such as attitude scales as well as cognitive measures.

Knapp (1973) points out that item sampling may be especially appropriate when the main purpose of testing is evaluation of a program rather than measurement of each individual relative to one another. Achievement monitoring (Gorth, Allen, Evans, O'Reilly, Pinsky, Wightman, 1970, Romberg, 1970) is a program evaluation model using item sampling. It is concerned with periodic testing within a course of study to help the teacher manage the instruction of individual students and revise the overall curriculum.

This study uses the achievement monitoring model to investigate one behaviorally-defined operationalization of the construct, resourcefulness.

Subjects

A sub-population of 24 students was drawn randomly from a population of approximately 500 students at the Interim Junior High School in Rochester, New York.

Interim was not organized by grades as in a traditional school. There were four subject matter areas--mathematics, science, social studies, and language arts. Each area had four or five teachers responsible for a group of about twenty students per period. The students spent one period in each area every day. There were also, at Interim, resource teachers for such activities as art, woodworking, communication, music, physical education, guidance. The students were expected to develop projects on which they work independently, or with selected others. The projects provided the basis for much of the learning which took place in school.

Items

Behavioral goals for the Interim Junior High School were continually being developed and revised. These goals were independent of specific content; they are directed to the principal focus of learning in the "open school", i.e., the acquisition and development of process skills.

The goal with which this study is concerned has been stated in two ways:

Institutional Objective - students will develop resourcefulness.

Individual Objective - the student will develop the ability to utilize environmental and internal resources.

Nine "items" were developed which related to this objective. These were derived from discussions with school personnel and from a search of the literature. An attempt was made to precisely define each item (Kunzelman, 1970).

Item 1. The number of different ways that a student's project shows creativity.

Rationale: An important aspect of learning at Interim is the development of independent activities, here defined as projects. Students engage in activities of interest, exploring new areas and reinforcing weak skills. Activities are problem oriented, involving the students with hands-on experiences using a wide range of materials. The activities may take place anywhere within the reach of the students, both within the schools and the local community (Learner-Centered School, 1971).

Torrance's creativity items (Torrance, 1967) are scored for flexibility which is defined as the number of different categories for response. The scoring adaptation of this item evaluates projects for the different ways in which a response (project content) is expressed. For example, projects ranged from a piece of creative writing (Score of 1) to an elaborate report on the "Jewish People" including drawings, written material and graphs (Score of 3).

Item 2. The number of different resources a student uses in his project.

Rationale: There is an attempt to provide a wide range of materials and many areas of interest at Interim. In addition, home and community are considered to be resource areas for students and teachers.

Teachers not only help students utilize resources but also encourage documentation of this process.

The Score for Item 2 was obtained by counting number of resources a student used in the development and completion of a project. This was obtained by a bibliography, if one was included in the project; or by direct questioning of the student. Each book, document, teacher, form of media was counted as 1 point toward total score.

Item 3. The number of "unusual uses" produced by a student.

"Most people throw their empty cardboard boxes away, but they have thousands of interesting and unusual uses. In the spaces below, and on the next page, list as many of these interesting and unusual uses as you can think of. Do not limit yourself to any one size of box. You may use as many boxes as you like. Do not limit yourself to the uses you have seen or heard about; think about as many possible new uses as you can." (Torrance, 1966, p. 35.)

Rationale: The definition of resourcefulness includes ability to utilize internal resources. Thinking of alternative responses would be a case of utilizing one's internal resources.

The fluency score, for the "unusual uses of tin cans" activity, is the number of different unusual uses produced. "Unusual use" is defined rather leniently as any relevant use other than the "usual use" as a container preserving food and other products. Other container responses detailing use after the can has been emptied of its original contents are counted. Fantastic or impossible uses are not counted.

Item 4. The number questions asked by a student during a one-half hour time period.

Rationale: Educators generally agree that asking questions in the school setting is crucial to productive learning. Children, in particular, seem to like to ask questions and to inhibit this tendency would seem to impede learning and constrict horizons.

Suskind's study (undated) of nine classrooms in a typical school setting showed a large number of teacher questions, a small number of student questions. He used observers in the classroom as recorders. He classified questions into Procedural, Recitational, Causal, Personal, Affective, Normative, and Other, and calculated separate and total scores.

For the present study the total number of questions asked were recorded. Observers were trained in a pilot study.

A school committed to the development of resourcefulness in youngsters would be expected to encourage question-asking behavior. In fact, informal discussions indicated that this is an area in which the teachers felt they have achieved some success.

Item 5. The number of extra-curricular activities.

Rationale: In an "open school" where curriculum is not traditional, extra-curricular needs its own definition. Youngsters have a planned program for the purposes of accountability. It includes conventional activites such as Math, and unconventional activities such as Yoga. Extra-curricular activities were defined as those not in planned program and included activities taking place both during school (running health food store, student government, "student teaching") and after school hours (athletic activities, music lessons, paper delivery).

A learner-centered approach which encourages youngsters to pursue their interests might be expected to have more youngsters participating in activities relevant to their needs and interests.

Item 6. The number of visits to the Communication Center during a one-week time period.

Rationale: The Communication Center at Interim consists of three areas-- a reference library, a typewriter and business skills area, and an audio-visual instructional media area.

Those who staffed the area viewed it as a place where youngsters could come for information-seeking and for gaining familiarity with audio and video equipment, typewriters. An open school encourages exploration of all sorts of educational materials.

Item 7. The number of times the student leaves the school building during the day for school related activities.

Rationale: An open school advocates utilization of resources in the community. Institutions such as libraries, museums, social agencies, as well as individuals--professional, service, etc., have been contacted and visited by youngsters in school.

Teachers have encouraged youngsters to individually, or in groups, seek out information outside of school, which would contribute to the development of a project.

Item 8. An unusual object was placed on the examiner's desk during the administration of Item 3. On each testing occasion a different object was used -- a collage composed of the inner workings of a clock, a ceramic Mexican bell, a group of origami figures. The number

of seconds required for the student to respond to the object were recorded. (A maximum time of 60 seconds was allowed.)

Rationale: An increased responsiveness to the environment would seem to be relevant to open education. More materials, more opportunity to interact with materials, and more opportunity to interact with other persons could be expected to increase a student's curiosity.

Curiosity, a responsiveness and openness to the environment, would seem to contribute to a youngster's ability to use environmental resources.

Torrance (1967) has suggested that a non-test way of measuring creativity is a person's response to a new situation. Curiosity has been studied in young children as a precursor to creativity. It has been defined as the ability to seek and maintain contact with novel stimuli. Mendel (1965) presented children with toy arrays, graduated with respect to novelty. Smock and Holt (1962) used perceptual conflict as a way of determining curiosity.

In this study, an "unusual object" was defined as something which would not be typically found on an examiner's desk. The object was placed on the desk, slightly off to one side, and paper and pencil were placed in front of the student. The examiner invited the student to sit down in front of the desk on which there was the paper and pencil and said, "I will be with you in a minute, I just have to find something in this folder." An observer was in the room while the examiner paused and then administered Item 3. Observer had a stop watch and noted response of student for 60 seconds. Observers were trained in the pilot study.

Item 9. The number of seconds the student spent on a task.

(Lite Brite*)

(Lite Brite consists of colored pegs which are stuck in a grid. Behind grid is a light. Designs can be made varying color and placement of pegs.)

Rationale: Perseverance, or the ability to stick with a task would seem to be positively related to resourcefulness and the stated goals of Interim.

An attempt was made to find a task which would have a universal appeal, not be sex-typed, and would depend minimally on intellectual capability. The Lite Brite seemed to meet these criteria.

*Hassenfeld Bros., Inc. Pawtucket, R.I., U.S.A. 1967.

The student was told "I am interested in how creative youngsters are at interim. I would like you to make a design, as creative or unusual as possible." A stop watch was used by the examiner to determine the amount of time spent.

Procedure

Administration of test items occurred three times during the second semester of the 1971-72 school year, at approximately one and one-half month intervals.

All items were administered to all students on each occasion. An item sampling a posteriori design was then imposed on the data. 100 overlapping samples of item data were drawn using a balanced incomplete block design, (Knapp, 1973) i.e.

From the total of nine test items, three were sampled for each student on each occasion. Each item appeared eight times in each design; each pair of items occurred two times in each design.

Item sampling seemed appropriate, even though the number of items was small, because of the time and resources involved collecting data for each item.

Analysis of Data

Each 24 x 9 population matrix was sampled on each testing occasion, and difference matrices between time 1 and time 2, time 2 and time 3, and time 1 and time 3 were also sampled.

The means and the standard deviations of the 100 samples on each occasion were calculated. Only the item means and standard deviations were analyzed; not the total test score.

These operations were performed by computer; an APL program was prepared for this purpose.

Results and Discussion

Twenty-eight students were administered nine items on the first testing occasion. Due to the transfer of students to other schools and lack of student cooperation, only twenty-five students were tested on the second occasion. There were twenty-four students still available for testing on the third occasion, but, some of the data were lost due to the unexpected early closing of the Rochester City Schools in June, 1972.

A summary of the principal findings for the twenty-four students for whom maximum data were available will focus on difference scores; since this area does not seem to have been previously explored.

Estimation of Mean Differences

Tables 1, 2, and 3 contain the difference scores for each subject on each item for each pair of occasions. Tables 4, 5, and 6 show a comparison of the population mean differences with the item sampling estimate of the mean differences, along with the standard deviations of the mean differences (standard error).

For item 1, the largest difference between the population parameter and the mean of the sample statistics if .07 (Table 6); it is .05 for item 2 (Table 5); .05 for item 3 (Table 4); .02 for item 4 (Table 6); .11 for item 5 (Table 6); .04 for item 6 (Table 4); .01 for item 7 (Table 4); 6 for item 8 (Table 4) and 194 for item 9 (Table 6).

Item sampling appears to give adequate estimates for items 1 - 8. The sampling mean difference deviates from the population mean difference in a range comparable to that of sample mean from the population mean. Item 9 had a wide spread of scores and the sample estimate is consequently more deviate from the population mean.

This completes the discussion of the results from the methodological point of view. Consideration of the substantive results now follow.

Monitoring Resourcefulness - Curriculum Evaluation

From the preceding data, some statements can be made about changes in resourcefulness of youngsters at Interim, as measured by items 1 - 9. Such changes, however, cannot be attributed solely to the curriculum since the design employed was evaluative rather than experimental.

Item 1. Flexibility, or number of modes of response, did not appear to develop noticeably during the semester. From Table 6 - the item mean increased across occasions by .08. One student showed a substantial gain as indicated in Table 3. In general, written reports were the most common mode of expression found in projects.

Item 2. The number of resources used in the development of projects increased slightly across occasions. The item mean gained .20 point as indicated in Table 6. Two students showed a substantial gain, one a substantial loss, as seen in Table 3. Utilization of a variety of resources by students grew minimally.

Item 3. The number of extra-curricular activities in which students participated increased over the semester. Table 6 indicated the item mean increased by .42. Three students showed an increase of two activities (Table 3). Activities in and out of school appealed to a greater number of youngsters as the semester progressed.

Item 4. The number of times students left the building during the school day does not appear to have increased substantially. The item mean increased by .13. Two students showed growth during the semester. Although more students leave the building for information seeking and other school-related activities than is typical in a conventional school, students at the Interim appear to spend most time in school.

Item 5. Use of the communications center increased slightly. The item mean increased by .71. Four students used the Center more frequently; one student used the Center much less across time. The lack of substantial gain on item scores may have been due to the lack of the Center's specified objectives. Many books and materials could be found in the areas, which meant that there was less need for students to use the Center.

Item 6. The number of questions asked by students has incomplete data for Occasion 3. Table 4 shows that the item mean gained .42 from Occasion 1 to Occasion 2. Two youngsters showed an increase in question-asking behavior by at least five questions, and two showed a decrease by at

least five questions, as indicated in Table 1. Closer inspection of the data shows a general trend in the direction of more question-asking benavior. This observation is confirmed by teachers at Interim, who noted in informal interviews that more questions were being asked about content, as well as about philosophy, motivation and purposefulness.

Item 7. Some data for Occasion 3 are missing on the Unusual Uses Item. Table 4 indicates a gain of .5833 from Occasion 1 to Occasion 2. Four students thought of a greater number of alternative uses, two students thought of fewer uses. The ability to use internal resources, as defined in this study, appears to be developing in youngsters at Interim.

Item 8. Some data for Occasion 3 is missing on the Unusual Object Item. Table 4 indicates a seconds gain of 17.71 from Occasion 1 to Occasion 2. Eight students took longer to notice the unusual object, four students took less time to notice the object as shown in Table 1. An increased responsiveness to the environment is not indicated by the above data as students generally noticed the unusual object with decreasing frequency. However, different objects were used on each occasion which may have had varying degrees of attractiveness for youngsters.

Item 9. The amount of time youngsters spent working on the Lite Brite decreased over time. The item mean across occasions decreased by 497.80 (Table 6). Five youngsters spent at least 1000 seconds less with the

Lite Brite as seen in Table 3. This item appeared to measure cooperativeness and docility rather than resourcefulness. Youngsters began to question the necessity of and rationale for this task and were less willing to spend time on it.

Conclusions

1. Item sampling appears to be an effective technique for monitoring unobtrusive measures.
2. Growth in resourcefulness of youngsters in the population as measured by "test" items has implications for curriculum assessment at the Interim Junior High School. Items which showed a negligible increase over time were Item 1 (Flexibility), Item 2 (Number of Resources) and Item 4 (Leaving School Building). Items which indicated growth in resourcefulness were Item 3 (Extra-Curricular Activities), Item 5 (Use of Communications Center), Item 6 (Number of Questions Asked), and Item 7 (Unusual Uses). Items which indicated a decrease in resourcefulness, as defined in this study, were Item 8 (Unusual Object) and Item 9 (Time on Lite Brite).

This study shows that item sampling can be useful for a wider range of data than achievement test scores or attitude scales. Measures such as the observations made in this study require much time and effort. Yet they are a helpful and a meaningful way to measure curricular developments in a school setting. Item sampling makes administration of such items

much more reasonable. Further investigations of this problem could include reliability studies, comparisons of item sampling with total population data using larger samples and additional demonstrations of item sampling's usefulness in behavioral, as well as cognitive, aspects of curriculum evaluation.

58:03:sf
(2/13/74)

BIBLIOGRAPHY

Bursack, B. and D. Cook. "Utilizing Item Sampling Techniques to Scale Affective Reactions to Mathematics." Paper presented at Annual Meeting of AERA, March 1970.

Gorth, W. P.; D. W. Allen; D. R. Evans; P. P. O'Rielly; P. D. Pinsky, and L. E. Wightman. Designing Instructional Systems with Longitudinal Testing Using Item Sampling Technique. A set of papers delivered at a Symposium at Annual Meeting of AERA, March 1970.

Knapp, T. R. "An Application of Balanced Incomplete Block Designs to the Estimation of Test Norms." Educational and Psychological Measurement, 28, 1968, pp. 265-272.

_____. Item-Examinee Sampling. Unpublished manuscript (University of Rochester, 1973).

Kunzelman, H. (Ed.). Precision Teaching. Seattle, Washington: Special Child Publications, 1970.

The Learner-Centered School, a Document Drafted in Planning Stages of Interim Junior High School, Rochester, New York, 1971.

Lord, F. M. "Estimating Norms by Item Sampling." Educational and Psychological Measurement, 22, 1962, pp. 259-267.

Mendel, Gisela. "Children's Preferences-for Differing Degrees of Novelty." Child Development, 36, 1965, pp. 443-465.

Petersen, D. F. and D. H. Anderson. "Closing the Communications Gap with Item Sampling." Paper presented at Annual Meeting of AERA, February, 1971.

Plumlee, Lynette B. "Estimating Means and Standard Deviations from Partial Data - An Empirical Check on Lord's Item Sampling Technique." Educational and Psychological Measurement, 24, 1969, pp. 623-630.

Pugh, R. C. "Empirical Evidence on the Application of Lord's Sampling Technique to Likert Items." Journal of Experimental Education, 39, 1971, pp. 54-56.

Romberg, T. A. "Achievement Monitoring via Item Sampling." Paper presented at the Annual Meeting of AERA, March, 1970.

Shoemaker, D. M. Principles and Procedures of Multiple Matrix Sampling. (in press) Inglewood, California: SWRL.

Sirotnik, K. "An Investigation of the Context Effect in Matrix Sampling." Journal of Educational Measurement, 7, 1970, pp. 199-208.

Smock, C. and Bess Holt. "Children's Reactions to Novelty: An Experimental Study of Curiosity Motivation." Child Development, 33, 1962, pp. 631-642.

Suskind, E. "The Role of Question Asking in the Elementary School Classroom" in The Psycho-Educational Clinic: Papers and Research, Edited by Sarason, Seymour, B. New York: Wiley (no date).

Torrance, E. P. Torrance Tests of Creative Thinking. Princeton, New Jersey: Personnel Press, Inc., 1966.

_____. "Non-Test Ways of Identifying the Creatively Gifted" in Creativity: Its Educational Implications, Edited by Gowan, Demos, and Torrance. New York: John Wiley and Sons, Inc., 1967.

166:03:sf
(2/13/74)

TABLES

TABLE 1

MATRIX W

POPULATION OF DIFFERENCE SCORES, $O_2 - O_1$

	ITEM								
EXAMINEE	1	2	3	4	5	6	7	8	9
1	-1	-1	0	-2	-1	4	-8	-10	30
2	0	1	-1	1	1	-4	-1	57	63
3	0	0	2	-1	-2	3	-1	0	77
4	1	0	1	-1	6	-5	1	-36	550
5	1	1	0	0	0	0	3	0	-757
6	0	1	0	4	0	-1	-2	14	-490
7	0	0	0	4	0	-4	2	35	-620
8	-1	0	3	0	-2	0	1	-19	-520
9	-1	0	1	3	0	5	0	-1	-837
10	0	0	1	1	-1	2	2	45	-40
11	-1	-1	0	0	-1	1	1	16	-228
12	1	2	3	1	4	3	0	28	-1040
13	0	0	0	-1	0	-3	2	0	-410
14	0	0	1	2	0	3	2	25	800
15	-1	0	0	0	0	1	1	46	380
16	1	1	0	-2	0	-3	3	-27	-2695
17	3	0	4	0	5	-6	-3	56	-170
18	-2	-4	-2	0	5	2	4	-5	-1065
19	0	0	1	3	7	2	1	60	-40
20	-1	0	2	2	0	1	4	4	-990
21	0	-1	0	-2	1	-1	1	10	614
22	-1	-2	0	-1	0	3	0	60	-510
23	1	1	3	-1	-4	-2	-1	60	60
24	-1	2	1	0	1	8	2	7	334

TABLE 2

MATRIX Y

POPULATION OF DIFFERENCE SCORES, $O_3 - O_2$

EXAMINEE	ITEM 1	2	3	4	5	6	7	8	9
1	0	1	0	0	0				-149
2	1	-1	2	-1	-1		-1	-56	-1771
3	1	0	0	0	1	-3	1	2	78
4	-1	0	-1	0	-3				-730
5	-1	-1	1	0	5	3	-3	0	-123
6	0	1	-1	-2	-1	-3			0
7	0	0	0	-1	0		1	0	-48
8	0	1	-1	-1	4		-1	-31	284
9	0	0	0	-2	0	-3	0	12	297
10	0	0	-1	1	0	2			74
11	0	2	2	2	0			-23	-132
12	0	1	-2	0	-4	-3	1	-28	53
13	0	1	0	0	2	3	1	-36	35
14	0	0	0	-3	1	-2	-1	10	-485
15	1	-1	1	0	0				-260
16	-1	-3	-1	0	3	5	1	47	300
17	-1	1	-4	-1	1	6	3	-39	-10
18	1	1	1	1	-1				-705
19	0	0	0	-1	-6	1	-2	0	-530
20	0	0	-1	-1	4	14	-2	4	-68
21	1	0	-1	0	-1	3	-1	-30	-284
22	2	1	0	1	0	-3	2	-30	405
23	-1	-1	-3	0	0	-2	1	-55	-30
24	1	1	0	0	-3	-3	3	-5	-74

TABLE 3

MATRIX Z

POPULATION OF DIFFERENCE SCORES, $O_3 - O_1$

	ITEM								
SUBJECT	1	2	3	4	5	6	7	8	9
1	-1	0	0	-2	-1				-119
2	1	0	1	0	0		2	-58	-1708
3	1	0	2	-1	-1	0	0	2	155
4	0	0	0	-1	0				-180
5	0	0	1	0	3	3	0	0	-870
6	0	2	-1	2	-1	-1			-490
7	0	0	0	3	0		3	33	-1200
8	-1	1	2	-1	2		2	-50	-238
9	-1	0	1	1	0	2	0	11	-540
10	0	0	-1	3	-1	4			-114
11	-1	1	2	2	-1		1	-7	-380
12	1	3	1	1	0	-2	1	0	-987
13	0	1	0	-1	2	0	3	-56	-375
14	0	0	1	-1	1	1	1	35	315
15	0	1	1	0	0				120
16	0	-2	-1	-2	3	2	4	20	-2395
17	2	1	0	-1	6	0	0	-3	-180
18	-1	-3	-1	1	4				-1770
19	0	0	1	2	1	3	-1	60	-570
20	-1	0	1	1	4	15	2	8	-1078
21	1	-1	-1	-2	0	2	0	-20	350
22	1	-1	0	0	0	0	2	10	-105
23	0	0	0	-1	-4	-4	0	5	30
24	0	3	1	0	-2	5	4	2	260

TABLE 4

ITEM SAMPLING RESULTS FOR DIFFERENCE SCORES
BETWEEN OCCASION 2 AND OCCASION 1
(MATRIX W)

ITEM	POPULATION MEAN DIFFERENCES	MEAN OF SAMPLE MEAN DIFFERENCES	STANDARD DEVIATION (STANDARD ERROR)
1	-0.0833	-0.0875	0.2949
2	0.0000	-0.0425	0.3698
3	0.8333	0.7800	0.4110
4	0.4167	0.4262	0.4395
5	0.7917	0.8337	0.7663
6	0.3750	0.4175	0.9791
7	0.5833	0.5900	0.6619
8	17.7083	11.6313	8.1189
9	-312.6667	-118.8943	208.4533

TABLE 5

ITEM SAMPLING RESULTS FOR DIFFERENCE SCORES
BETWEEN OCCASION 3 AND OCCASION 2
(MATRIX Y)

ITEM	POPULATION MEAN DIFFERENCES	MEAN OF SAMPLE MEAN DIFFERENCES	STANDARD DEVIATION (STANDARD ERROR)
1	0.1250	0.1687	0.2783
2	0.2500	0.2025	0.2705
3	-0.4167	-0.4012	0.4153
4	-0.2917	-0.3012	0.3487
5	-0.0833	-0.0900	0.8508
6			
7			
8			
9	-185.1250	-188.7710	136.8896

TABLE 6

ITEM SAMPLING RESULTS FOR DIFFERENCE SCORES
BETWEEN OCCASION 3 AND OCCASION 1
(MATRIX Z)

ITEM	POPULATION MEAN DIFFERENCES	MEAN OF SAMPLE MEAN DIFFERENCES	STANDARD DEVIATION (STANDARD ERROR)
1	0.0417	0.1100	0.2318
2	0.2500	0.2387	0.3594
3	0.4167	0.4087	0.2825
4	0.1250	0.1487	0.4581
5	0.7083	0.5938	0.6880
6			
7			
8			
9	-497.7917	-529.9907	206.6935